Massachusetts Institute of Technology
Department of Economics
Working Paper Series

IV ESTIMATION WITH VALID AND INVALID INSTRUMENTS

Jinyong Hahn
Jerry Hausman

Room E52-251
50 Memorial Drive
Cambridge, MA 02142
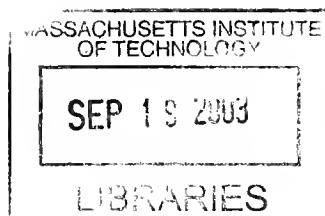
IV Estimation with Valid and Invalid Instruments

Jinyong Hahn and Jerry Hausman
UCLA and MIT

Abstract

While 2SLS is the most widely used estimator for simultaneous equation models, OLS may do better in finite samples. Here we demonstrate analytically that the for the widely used simultaneous equation model with one jointly endogenous variable and valid instruments, 2SLS has smaller MSE error, up to second order, than OLS unless the R^2, or the F statistic of the reduced form equation is extremely low. We then consider the relative estimators when the instruments are invalid, i.e. the instruments are correlated with the stochastic disturbance. Here, both 2SLS and OLS are biased in finite samples and inconsistent. We investigate conditions under which the approximate finite sample bias or the MSE of 2SLS is smaller than the corresponding statistics for the OLS estimator. We again find that 2SLS does better than OLS under a wide range of conditions. We then present a method of sensitivity analysis, which calculates the maximal asymptotic bias of 2SLS under small violations of the exclusion restrictions. For a given correlation between invalid instruments and the error term, we derive the maximal asymptotic bias. We apply our results to IV estimation of the returns to education. We derive the bias in the estimated standard errors of 2SLS for the first time. This derivation also has implications for the test of over-identifying restrictions.

IV Estimation with Valid and Invalid Instruments

Jinyong Hahn and Jerry Hausman
July 15, 2003

While 2SLS is the most widely used estimator for simultaneous equation models, OLS may do better in finite samples. Econometricians have recognized this possibility, and many Monte Carlo studies were undertaken in the early years of econometrics to attempt to determine condition when OLS might do better than 2SLS. Here we demonstrate analytically that the for the widely used simultaneous equation model with one jointly endogenous variable and valid instruments, 2SLS has smaller MSE error, up to second order, than OLS unless the R^2, or the F statistic of the reduced form equation is extremely low. We do a calculation based on observable statistics with one unknown parameter that allows a calculation that should give valuable information about the relative MSEs of OLS and 2SLS.

We then consider the relative estimators when the instruments are invalid, i.e. the instruments are correlated with the stochastic disturbance. Here, both 2SLS and OLS are biased in finite samples and inconsistent. We investigate conditions under which the approximate finite sample bias or the MSE of 2SLS is smaller than the corresponding statistics for the OLS estimator. We again find that 2SLS does better than OLS under a wide range of conditions, which we characterize as functions of observable statistics and one unobservable statistic.

We then present a method of sensitivity analysis, which calculates the maximal asymptotic bias of 2SLS under small violations of the exclusion restrictions. For a given correlation between invalid instruments and the error term, we derive the maximal asymptotic bias. We demonstrate how such maximal asymptotic bias can be estimated in practice.

Next, we turn to inference. In the "weak instruments" situation the bias in the 2SLS estimator creates a problem, since it is biased towards the OLS estimator, which is also biased. The other problem that arises is that the estimated standard errors of the 2SLS estimator are often much too small to signal the problem of imprecise estimates.

1

Here we derive the bias in the estimated standard errors for the first time, which turns out to cause the problem. This derivation also has implications for the test of over-identifying restrictions.

We do not survey the weak instruments literature. For recent surveys see Stock et. al. (2002) and Hahn and Hausman (2003).

I   Model Specification

We begin with the model specification with one right hand side (RHS) jointly endogenous variable so that the left hand side (LHS) variable depends only on the single jointly endogenous RHS variable. This model specification accounts for other RHS predetermined (or exogenous) variables, which have been "partialled out" of the specification. We will assume that[1]

(1.1) $y_1 = \beta y_2 + \varepsilon_1$

(1.2) $y_2 = z\pi_2 + v_2,$

where $\dim(\pi_2) = K$. Thus, the matrix z is the matrix of all predetermined variables, and equation (1.1) is the reduced form equation for y_2 with coefficient vector π_2. We also assume homoscedasticity:

(1.3) $\begin{pmatrix} \varepsilon_{1i} \\ v_{2i} \end{pmatrix} \sim N(0, \Sigma) \sim N\left(0, \begin{bmatrix} \sigma_{\varepsilon\varepsilon} & \sigma_{\varepsilon v} \\ \sigma_{\varepsilon v} & \sigma_{vv} \end{bmatrix}\right).$

We use the following notation:

$y \equiv \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad z \equiv \begin{pmatrix} z_1' \\ \vdots \\ z_n' \end{pmatrix}, \quad \sigma_{\varepsilon\varepsilon} \equiv \mathrm{Var}(\varepsilon_{1i}), \ \sigma_{vv} = Var(v_{2i}), \quad \sigma_{\varepsilon v} \equiv \mathrm{Cov}(\varepsilon_{1i}, v_{2i}) = \sigma_{12}.$

We initially assume the presence of valid instruments, $E[z'\varepsilon / n] = 0$ and $\pi_2 \neq 0$.

---

[1] Without loss of generality we normalize the data such that y_2 has zero mean.

II   Estimation with Valid Instruments

From previous papers, e.g. Hahn and Hausman (2002a, 2002b) we know the bias and MSE of 2SLS up to second order. The bias of 2SLS is

$$(2.1) \quad E[b_{2SLS}] - \beta \approx \frac{K\sigma_{\epsilon v}}{n\Theta} \approx \frac{K\sigma_{\epsilon v}}{nR^2 \operatorname{var}(y_2)},$$

where $\Theta = \pi'z'z\pi/n$, assumed to be fixed, R^2 is the theoretical value from the second (reduced form) equation, and y_2 is normalized to have mean zero. We assume:

Condition 1: $K \to \infty$ as $n \to \infty$ such that $K/\sqrt{n} = \mu + o(1)$ for some $\mu \neq 0$.

A   Properties of the 2SLS Estimator

As a special case of Theorem 3 in Section 3, we obtain that:

Theorem 1: $\sqrt{n}(b_{2SLS} - \beta) \Rightarrow N\left(\frac{\sigma_{\epsilon v}\mu}{\Theta}, V_{2SLS} \right).$

Here, $V_{2SLS} = \sigma_{\epsilon\epsilon}/\Theta$, the usual 2SLS first order asymptotic variance. As a consequence, we obtain the approximate MSE of 2SLS:

$$(2.2) \quad MSE(2SLS) = M_2 \approx \frac{\sigma_{\epsilon v}^2 \mu^2}{n\Theta^2} + \frac{V_2}{n} \approx \frac{K^2 \sigma_{\epsilon v}^2}{R^4 (y'_2 y_2)^2} + \frac{\sigma_{\epsilon\epsilon}}{R^2 (y'_2 y_2)}$$

Note that both terms in equation (2.2) approach zero as $(y_2'y_2)$ increases with increasing sample size. The first term, bias squared also approach zero more quickly, as expected, since 2SLS is "root n" consistent.

We now simplify the M_2 expression for 2SLS. Without loss of generality we use the normalization (rescaling of units) $\sigma_{\epsilon\epsilon} = \sigma_{vv} = 1$ so that $Var(y_2) = 1/(1-R^2)$ and $\sigma_{\epsilon v} = \rho.$[2] Using this normalization we find:

$$(2.3) \quad MSE(2SLS) = M_2 \approx \left(\frac{K^2 \rho^2 (1-R^2) + nR^2}{n^2 R^2} \right)\left(\frac{1-R^2}{R^2} \right)$$

The convergence of MSE to zero in terms of the sample size n becomes quite evident with this normalization.

---

[2] These parameter are theoretical values from the underlying model specifications for given parameter values.

B Properties of the OLS Estimator

We now calculate the bias and MSE of the OLS estimator. The approximate bias is:

$$(2.4) \quad E[b_{OLS}] - \beta \approx \frac{\mathrm{Cov}(y_2, \varepsilon)}{\mathrm{Var}(y_2)} \approx \frac{\sigma_{\varepsilon v}}{\Theta + \sigma_{vv}}$$

The approximate variance is defined as:

$$(2.5) \quad V_{OLS} = \frac{\sigma_{\varepsilon\varepsilon}}{\Theta + \sigma_{vv}} - \frac{\sigma^2_{\varepsilon v}}{(\Theta + \sigma_{vv})^2} - \frac{2\sigma^2_{\varepsilon v}\Theta^2}{(\Theta + \sigma_{vv})^4}$$

As a special case of Theorem 4 in Section 3, we obtain the distribution for the OLS estimator:

Theorem 2: $\sqrt{n}\left(b_{OLS} - \left(\beta + \frac{\sigma_{\varepsilon v}}{\Theta + \sigma_{vv}} \right) \right) \to N(0, V_{OLS})$

This result is the same as in Hausman (1978). Thus, the approximate MSE of OLS is

$$(2.6) \quad MSE(OLS) = M_0 \approx \frac{\sigma^2_{\varepsilon v}}{(\Theta + \sigma_{vv})^2} + \frac{V_{OLS}}{n}$$

The inconsistency of OLS is evident from equation (2.6) because while the second term goes to zero as n becomes large, the first term is not a function of *n*. The OLS MSE under the normalization used above becomes:

$$(2.7) \quad \begin{aligned} M_0 &\approx \frac{\sigma^2_{\varepsilon v}}{\mathrm{var}^2(y_2)} + \frac{\sigma_{\varepsilon\varepsilon}}{n\,\mathrm{var}(y_2)} - \frac{\sigma^2_{\varepsilon v}}{n\,\mathrm{var}^2(y_2)} - \frac{2\sigma^2_{\varepsilon v}R^4}{n\,\mathrm{var}^2(y_2)} \\ &= \left(\frac{\rho^2(n - 1 - 2R^4)(1 - R^2) + 1}{n} \right)(1 - R^2) \end{aligned}$$

We see that the first term in the OLS MSE is the usual first order bias term squared. It does not go to zero as n becomes large since OLS is inconsistent.

From the 2SLS MSE calculations as the sample size grows large the denominator of the 2SLS MSE calculation in equation (2.3) dominates, and the MSE goes to zero. To the contrary for the OLS MSE, in equation (2.7) the numerator also grows with n, as the bias of the OLS estimator does not go to zero with the sample size. Thus, for large samples 2SLS is consistent and OLS is not. We now consider how the estimators do in finite samples.

C   Bias Comparisons of the 2SLS and OLS Estimators

We compare the approximate finite sample bias of 2SLS to the approximate MSE of OLS:

$$(2.8) \quad \frac{B_2}{B_0} = \frac{K}{nR^2} = \frac{1}{1-R^2} \frac{1}{F}$$

where the F statistic is the "theoretical" F-statistic from the first-stage reduced form. Thus, if $F >> 1$, 2SLS has less bias. However the OLS variance is less than the 2SLS variance so we compare the MSEs below.

Before leaving the bias comparisons, we also consider what happens when we are close to being unidentified so that $\pi_2 = a/\sqrt{n}$, where the vector has dimension K. Thus, the reduced form coefficients are "local to zero". With $\pi_2 = a/\sqrt{n}$, equation (2.1) predicts the bias of 2SLS to be

$$(2.9) \quad E[b_{2SLS}] - \beta \approx \frac{K\sigma_{\varepsilon v_2}}{\Psi} = \frac{\sigma_{\varepsilon v_2}}{\frac{1}{K}\Psi}$$

where equation (2.9) is an approximation to the asymptotic bias of 2SLS under the asymptotics where $\pi_2 = a/\sqrt{n}$. Here, $\Psi = a'z'za$. On the other hand, equation (2.4) predicts the approximate bias for OLS to be:

$$(2.10) \quad E[b_{OLS}] - \beta \approx \frac{\sigma_{\varepsilon\upsilon}}{\frac{1}{n}\Psi + \sigma_{vv}}$$

Taking the ratio of the biases under local to zero asymptotics:

$$(2.11) \quad \frac{B_{L2}}{B_{LO}} = \frac{\frac{1}{n}\Psi + \sigma_{vv}}{\frac{1}{K}\Psi}$$

From equation (2.11), it follows that the bias of 2SLS is smaller than OLS as long as $K \ll n$, a condition which will always be satisfied in practice.

D   MSE Comparisons of the 2SLS and OLS Estimators

We next compare the MSE of 2SLS to the MSE of OLS using the normalization (and non-local asymptotics):

$$(2.12) \quad \frac{M_2}{M_0} = \frac{K^2\rho^2(1-R^2) + nR^2}{(nR^4)[(n-1-2R^4)\rho^2(1-R^2)+1]}$$

where $R^4 = (R^2)^2$. The correlation parameter ρ is the key parameter in simultaneous equation analysis because if it is zero the OLS estimator is the unbiased Gauss-Markov estimator and the ratio of MSEs in equation (2.12) equals $1/R^2 > 1$, but OLS is biased and inconsistent if the parameter value of ρ is not zero.

Which estimator to use will depend on whether equation (2.12) is less than or greater than unity. We can solve for the "critical value" of ρ^2 which causes the MSE of the 2 estimators to be equal. The solution for this "critical value" has a remarkably simple form:

$$(2.13) \quad \rho^2 = \frac{nR^2}{nR^4(n-1-2R^4) - K^2}$$

As n becomes large the "critical value" of ρ^2 goes to zero. In any particular sample R^2 and F can typically be accurately estimated from the unbiased estimates of the reduced form so that only ρ^2 is unknown. While this parameter value is typically unknown, the applied econometrician will often have a good (a priori) knowledge of ρ so that she will be able to determine whether the critical value is below the square of the correlation coefficient.[3] As we now demonstrate, the critical value is often so low that 2SLS will have a lower MSE than OLS, even for situation with relatively "weak instruments" or a low F statistic.

In Figure 1 we calculate the critical value of ρ (using the absolute value) for a range of values of R^2 for K of 5, 10, and 30 and for sample sizes of n = 100.[4] The results of Figure 1 demonstrate that for K=5 if $R^2 \geq 0.1$ then the critical value of ρ is sufficiently small that 2SLS should typically be used in terms of the MSE comparison. For K=10 2SLS will typically be better if $R^2 \geq 0.2$. However, for K=30 we typically require $R^2 \geq 0.4$. In Table 1 we repeat the calculations for n=500 and n=1000. Here we find that if $R^2 \geq 0.1$ that 2SLS typically will have a lower MSE. Thus, except in the case of weak instruments, which can arise when both R^2 is low and the number of instruments is high, 2SLS is typically the preferred estimator based on an approximate finite sample comparison of MSEs.

III   Estimation with Invalid Instruments

Up to this point we have assumed that the instruments are valid so that they are orthogonal to the stochastic disturbance ε_1. However, the econometrician may not be certain that the instruments satisfy the orthogonality condition. We now consider the situation where the orthogonality condition on the instruments fails so that $E[z'\varepsilon_1 / n] \neq 0$. We first consider the "large sample bias" of 2SLS:

---

[3] The parameter ρ is also estimated from the 2SLS estimation, but a good estimate may be diffcult to achieve in a "weak instrument" situation.
[4] The curves for increasing K lie to the right of each other.

$$(3.1) \quad \text{plim}[b_{2SLS}] - \beta \approx \frac{\sigma_{W'\varepsilon}}{R^2 \sigma_{y_2 y_2}}$$

where $W = z\pi_2$. When we compare this with the analogous expression for OLS

$$(3.2) \quad \text{plim}[b_{OLS}] - \beta \approx \frac{\sigma_{y_2'\varepsilon}}{\sigma_{y_2 y_2}}$$

In general either estimator may be preferred on this criterion depending on circumstances. The numerator of equation (3.1) would likely be smaller ("less correlation" in the instrument) than the numerator of equation (3.2), but the denominator of equation (3.1) is always smaller since $R^2 < 1$. Indeed, if R^2 is very small, the OLS estimator may do better in terms of inconsistency.

A   Invalid Instrument Specification

To do asymptotic approximations we need to specify the correlation of the instrument with the stochastic disturbance in the structural equation (1.1). We use a local specification similar to the approach in Hausman (1978, Theorem 2.1):

$$(3.3) \quad \varepsilon_1 = z(\gamma / \sqrt{n}) + e \; \textit{ for } \gamma \neq 0.$$

We assume that (e, v) is homoscedastic and zero mean normally distributed with covariance matrix :

$$\begin{pmatrix} e_{1i} \\ v_{2i} \end{pmatrix} \sim N(0, \Omega) \sim N\left(0, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \right)$$

B   Properties of the 2SLS Estimator with Invalid Instruments

We derive the asymptotic distribution of the 2SLS estimator with locally invalid instruments in Appendix A:

Theorem 3: $\sqrt{n}(b_{2SLS} - \beta) \Rightarrow N\left(\dfrac{\Xi + \mu\sigma_{12}}{\Theta}, V_{2SLS} \right) \approx N\left(\dfrac{\sqrt{n}\sigma_{W'\varepsilon} + \left(K/\sqrt{n}\right)\sigma_{12}}{R^2\sigma_{vv}}, V_{2SLS} \right)$

where $W = z\pi_2$ is the instrument and $\Xi = \pi'z'z\gamma/n$, which is assumed to be fixed . The first term in the numerator of the mean Ξ arises from failure of the orthogonality condition. The second term is the usual finite sample bias term and it decreases with the sample size. The variance continues to be V_{2SLS} under instrument invalidity because of the local departure in equation (3.3) similar to Hausman (1978, p. 1256).

We use Theorem 3 to calculate the approximate bias of the 2SLS estimator with invalid instruments is:

$$(3.4) \quad B_{2I} \approx \frac{\Xi/\sqrt{n} + K\sigma_{12}/n}{\Theta} = \frac{1-R^2}{R^2}\left(\frac{1}{\sqrt{n}}\alpha\rho + \frac{1}{n}K\rho \right),$$

where we use the previous normalizations and set $\sigma_{W'\varepsilon} = \Xi/\sqrt{n} = \alpha\rho/\sqrt{n}$ for $\alpha < 1$. Using Theorem 3 we find the MSE of 2SLS to be:

$$(3.5) \quad \begin{aligned} MSE_{2I} &\approx \frac{(\Xi + \mu\sigma_{12})^2}{\Theta^2}\frac{1}{n} + \frac{V_{2SLS}}{n} \\ &= \left(\frac{1-R^2}{R^2} \right)^2 \left(\frac{1}{\sqrt{n}}\alpha\rho + \frac{1}{n}K\rho \right)^2 + \frac{1}{n}\left(\frac{1-R^2}{R^2} \right) \end{aligned}$$

C Distribution of the OLS Estimator with Invalid Instruments

We derive the asymptotic distribution of the OLS estimator with locally invalid instruments in Appendix A:

Theorem 4: $\sqrt{n}\left(b_{OLS} - \left(\beta + \dfrac{\sigma_{12}}{\Theta + \sigma_{22}} \right) \right) \Rightarrow N\left(\dfrac{\Xi}{\Theta + \sigma_{22}}, V_{OLS} \right) \approx N\left(\dfrac{\sqrt{n}\sigma_{W'\varepsilon}}{\sigma_{vv}}, V_{OLS} \right)$

The distribution is centered around the usual OLS bias, as before, and the numerator of the mean of the distribution arises from the instrument invalidity. Again, the variance

9

continues to be V_{OLS} under instrument invalidity because of the local departure in equation (3.3). Using Theorem 4 we find the MSE of OLS to be:

$$(3.6) \quad \begin{aligned} MSE_{OI} &\approx \left(\frac{\sigma_{12}}{\Theta + \sigma_{22}} + \frac{1}{\sqrt{n}} \frac{\Xi + \mu\sigma_{12}}{\Theta + \sigma_{22}} \right)^2 + \frac{V_{OLS}}{n} \\ &= (1 - R^2)^2 \left(\frac{1}{\sqrt{n}} \alpha\rho + \rho \right)^2 + \left(\frac{-\rho^2(1 + 2R^4)(1 - R^2) + 1}{n} \right)(1 - R^2) \end{aligned}$$

The first term in parentheses is the "usual" simultaneous equation bias of OLS that does not decrease with the sample size.

We consider a special situation which make the formulae easier to interpret. Let $\gamma = \tau\pi$ for some τ. Under this proportionality assumption, the asymptotic distributions take the form:

$$\sqrt{n}(b_{2SLS} - \beta) \Rightarrow N\left(\tau + \frac{\mu\sigma_{12}}{\Theta}, V_{2SLS} \right) \qquad \text{and}$$

$$\sqrt{n}(b_{OLS} - (\beta + \frac{\sigma_{12}}{\Theta + \sigma_{22}}) \Rightarrow N\left(\frac{\tau}{1 + \sigma_{22}/\Theta}, V_{OLS} \right) \approx N\left(\tau R^2, V_{OLS} \right)$$

where we have used the normalization to derive the final expression for the distribution of OLS.

D   Bias Comparison of 2SLS with OLS

We now compare the bias of 2SLS under instrument invalidity with the bias of OLS given similar circumstances. We now re-write the bias of OLS using the normalization:

$$(3.7) \quad B_{OI} = E[b_{OLS}] - \beta \approx (1 - R^2)\left(\frac{1}{\sqrt{n}} \alpha\rho + \rho \right)$$

As before, we take the ratio of (3.4) and (3.7):

$$(3.8) \quad \frac{B_{2I}}{B_{OI}} = \frac{\alpha\rho/\sqrt{n} + K\rho/n}{R^2(\alpha\rho/\sqrt{n} + \rho)}.$$

The ratio of the biases is homogeneous of degree zero in the correlation coefficient ρ, so we can simplify terms. We plot the ratio of the biases in Figure 2 for the case of n=100 and K=5 and $\alpha = 0.1$.

We find that the 2SLS bias is less than the OLS bias if:

$$(3.9) \quad \alpha < \frac{1}{\sqrt{n}} \frac{nR^2 - K}{1 - R^2}.$$

Equation (3.9) is very easy to interpret. We calculate a "critical alpha" in Figure 3, and note that it increase quite rapidly, so that the bias of 2SLS with invalid instruments remains less than the bias of OLS so long as F exceeds 1.0 by a small amount. The straightforward relationship of equation (3.9) allows for an easy interpretation on which the econometrician may well have some a priori knowledge.

In certain situations it may be reasonable to consider a relationship between $\sigma_{w'\varepsilon}$ and R^2 such that when the covariance is less so is R^2. If we totally differentiate equation (3.9), we find that $dR^2/d\alpha = (n-K)/\left(n^{1/2}\left(\alpha + \sqrt{n}\right)^2\right)$. Thus, for a given increase in the covariance between the instrument and the stochastic term, α, we find that the required increase in R^2 is approximate at the rate of the one over the square root of n to keep the ratio of the biases approximately the same. However, the required change in R^2 is also inversely related to α.

Note that the common empirical finding that the 2SLS coefficient is larger than the OLS coefficient can arise because of the OLS bias when the instruments are valid or because of an improper instrument. Thus, even if the instrument is "almost uncorrelated" so that $\sigma_{w'\varepsilon} \approx 0$ substantial bias can still arise because R^2 is often quite small in the weak instruments situation. Thus, comparing equation (3.4) to the bias of OLS in equation (3.7), the empirical finding that the 2SLS estimate increases compared to the OLS estimate may indicate that the instrument is not orthogonal to the stochastic

disturbance. The resulting bias can be substantial. Indeed, it could exceed the OLS bias, leading to an increase in the estimated 2SLS coefficient over the estimated OLS coefficient.

E   MSE Comparison of 2SLS and OLS with Invalid Instruments

Returning to the general situation and using the normalizations the ratio of the MSEs is

$$(3.10) \quad \frac{M_2}{M_O} = \frac{(1 - R^2)(\alpha\rho + K\rho/\sqrt{n})^2 / R^4 + 1/R^2}{(1 - R^2)[(\alpha\rho + \sqrt{n}\rho)^2 + 1 - (1 - R^2)\rho^2 - 2(1 - R^2)\rho^2 R^4]} .$$

No straightforward condition can be derived where the ratio is less than one. We graph the ratio of the MSEs for $\alpha = 0.1$ and K=5, n=100 in Figure 4 (please note the inverted vertical axis). Note that the ratio of MSEs is below 1.0 except in the situation where R^2 becomes quite small (as with weak instruments) **and** ρ becomes small (which decreases the OLS bias). The situation remains essentially the same when we increase to $\alpha = 0.3$ in Figure 5. To yield a better understanding of what can happen in this situation, we plot the situation in Figure 6 where $R^2 \leq 0.2$ *and* $\rho \leq 0.3$ for $\alpha = 0.1$. Figure 6 demonstrates that the 2SLS estimator can do quite poorly compared to the OLS estimator, even though the F statistic exceeds 1.0 by a large amount. The reason for this poor relative performance is the small size of ρ which makes OLS a relatively good estimator. However, this situation is typically not a situation where the absolute performance of the 2SLS estimator with valid instruments would be poor under weak instruments because ρ is not large. It is the presence of invalid instruments, with only a "small amount" of correlation with the stochastic disturbance that creates the problem.

F.   Comparison of a (Second order) Unbiased Estimator

In our comparisons of 2SLS with OLS, two sources of bias arise. The first source of bias is from the use of estimated parameters, $\hat{\pi}_2$ in equation (1.2), in forming the instruments. This source of bias disappears as the sample becomes large. The second source of bias is from the use of invalid instruments, $\gamma \neq 0$ in equation (3.3). This source

12

of bias does not disappear sufficiently fast with the sample size to cause 2SLS to be consistent. An interesting question would be about how the comparison of IV to OLS would change if the first source of bias were eliminated. We can eliminate this source of bias (to second order) by using the Nagar estimator.[5]

We derive the asymptotic distribution of the Nagar estimator with locally invalid instruments in Appendix A:

Theorem 5: $\sqrt{n}(b_N - \beta) \Rightarrow N\left(\dfrac{\Xi}{\Theta}, \dfrac{\sigma_{\varepsilon\varepsilon}}{\Theta}\right) \approx N\left(\dfrac{\sqrt{n}\sigma_{W'\varepsilon}}{R^2\sigma_{yy}}, V_{2SLS}\right)$

where $W = z\pi_2$ is the instrument and $\Xi = \pi'z'z\gamma/n$, which is assumed to be fixed, and as before $V_{2SLS} = \sigma_{\varepsilon\varepsilon}/\Theta$. Thus to compare the MSE of the Nagar estimator to the MSE of the 2SLS estimator with invalid instruments, we see that the variance of the two estimators is the same, but that the bias differs as explained above. However, when we compare the bias square of 2SLS from equation (3.4) with the Nagar estimator we find that

$$(3.11) \quad \left(\frac{\Xi/\sqrt{n} + K\sigma_{12}/n}{\Theta}\right)^2 - \left(\frac{\Xi/\sqrt{n}}{\Theta}\right)^2 = \left(\frac{2K\sigma_{12}\Xi/n^{3/2} + K^2\sigma_{12}{}^2/n^2}{\Theta^2}\right)$$

can be less than or greater than zero. Thus, we cannot conclude that using the Nagar estimator to compare with OLS would make the comparison more favorable to an IV estimator.

IV   Sensitivity Analysis

Card (2001) discusses possible concerns that the instruments may be invalid in discussing the empirical literature that estimates the return to additional education. The use of instrumental variables in this situation began with Griliches (1977) well known paper. To investigate the possibility of invalid instruments, we consider the specification:

---

[5] The Nagar estimator may perform poorly with weak instruments because of its lack of moments. See Hahn, Hausman, and Kuersteiner (2002).

$$(5.1) \quad y_1(\theta) = \beta y_2 + z\theta + \varepsilon^*$$
$$\varepsilon^* = z\theta + \varepsilon$$

Note that we have added $z\theta$ to the error ε.[6] We derive the maximal asymptotic bias for a small violation of the exclusion restriction in Appendix B, where ψ is the correlation between $z_i\pi$ and ε_i^* so that ψ^2 is the R^2 of between $z_i\pi$ and ε_i^*. We find the maximal asymptotic bias to be:

Theorem 6: $\quad \max\left|\text{plim } bias \ \hat{\beta}_{2SLS}(\theta)\right| = \left(\frac{1}{R^2}\frac{\text{plim}n^{-1}\sum \varepsilon_i^2}{\text{plim}n^{-1}\sum y_{2i}^2}\right)^{1/2}\left(\frac{\psi^2}{1-\psi^2}\right)^{1/2}$

Note that the maximal asymptotic bias can be consistently estimated by

$$(5.2) \quad \left(\frac{1}{R^2}\frac{n^{-1}\sum \hat{\varepsilon}_i^2}{n^{-1}\sum y_{2i}^2}\right)^{1/2}\left(\frac{\psi^2}{1-\psi^2}\right)^{1/2}$$

Imbens (2003) suggested a different sensitivity analysis in a program evaluation model with binary explanatory variable, extending Rosenbaum and Rubin's (1983). With some simplification, it can be said that he considers a parametric model where an omitted variable bias is suspected. It is well known that the omitted variable bias can be related to two parameters, the coefficient of the omitted variable and the correlation of the omitted variable with other observed variables in the model, e.g. Griliches (1957). The sensitivity analysis of Imbens (2002) is based on manipulation of these two parameters.

We now consider the effect of invalid instruments in an empirical example. Estimating the return to education has been a well-researched problem over the past 25 years. Griliches (1977) is a seminal paper that uses IV to estimate returns to schooling. The usual result is that researchers find the OLS estimate to be smaller than the 2SLS estimate by approximately 25%-50%, e.g. Card (2001). This result arises from a tradeoff between two potential sources of bias: (1) an omitted variable, call it "spunk" in the stochastic disturbance may be correlated with the amount of educations. Thus, people

---

[6] Imbens (2003) considers the question of sensitivity analysis, but not in the context of instrumental variables.

with more spunk achieve higher education levels and also higher earnings, because they work harder both in school and on the job. This left out variable would lead to an upward bias in the least squares estimate of the schooling coefficient. (2) errors in variables (EIV) that arise because years of schooling are a noisy measure of "useful knowledge" attained with more years of school that leads to higher earnings. Here the EIV would lead to a downward bias in the least squares estimate of the schooling coefficient. The typical IV results finds that the EIV effect is larger than the left out variable effect, so the 2SLS estimated typically exceed the OLS results by a significant amount.

We now consider the Angrist and Krueger (AK) results which have an extremely large sample and use quarter of birth to form instruments, which may be more likely to be orthogonal to the stochastic disturbance than more widely used family background and other types of instruments typically used in the empirical returns to schooling literature. However, the AK instruments have an extremely low R^2 that could help create a weak instruments situation. Angrist and Krueger (1991) used a sample of n = 329,509 observations to estimate the returns to education. Using the AK data we estimate the 2SLS return to education to be 0.0891 (.016) using K=30 after partialing out the other right hand side variables. This estimate is closer to the OLS estimate of 0.071 (.0003) than expected given other empirical results. After partialling out, we find that the average squared residuals equal 0.41, the average of the partialled out right hand side endogenous variable (education) equals 10.8, and R^2 = .00044662. For $\psi^2 = 0.0001$, we find that the solution to equation (5.2) is 0.0925. This maximal bias exceeds the 2SLS estimate of 0.0891, so a small amount of bias could either eliminate any estimated return to education or double the estimate.

Our finding that the returns to coefficient could be over two times the OLS estimate contrasts with the results of Manski and Pepper (2000) who apply Manski's (1990, 2003) non-parametric bounds approach. Manski and Pepper (on a different sample) find that the upper bound is substantially less than two times the usual OLS estimates of the returns to schooling. However, the Manski approach does not allow for

15

errors in variables. This omission may significantly limit the empirical relevance of the Manski approach to this problem.[7]

This result demonstrates that use of a "weak identification" strategy such as the AK approach is extremely sensitive to very small departures from the IV orthogonality assumption. Note that from the result in Theorem 6, that this extremely sensitivity does not decrease with increasing sample size. Thus, the AK estimate of the returns to schooling is very sensitive despite an extremely large sample size of $n = 329,509$. Our results caution against using a weak identification strategy that has become widely used in applied econometrics.

V   Bias in Estimated Standard Errors

We have previously discussed the biased in the 2SLS estimator in equation (2.1) and Theorem 1. In the "weak instruments" situation this bias may be quite large. A further problem arises in that the 2SLS estimator is biased in the same direction as the OLS estimator as equation (2.4) and Theorem 2 demonstrate. Thus, Hausman (1978) specification type test will be biased towards not rejecting the null hypothesis of lack of orthogonality between ε_1 and v_2 in equations (1.1) and (1.2). However, another problem has been recognized in the weak instruments situation. The estimated standard errors for the 2SLS estimator are downward biased, sometimes leading to the mistaken inference that the 2SLS estimate are much more precise than they actually are. From analysis based on first order asymptotics the usual conclusion would be that with "weak instruments" that the reported standard error of the 2SLS estimator would be sufficiently large to signal the finding that so much uncertainty exists with the estimate that it would not be of much use. However, researchers have found that, to the contrary, often the 2SLS estimator in the presence of weak instruments leads to a reasonably small standard error. Thus, the researcher may be unaware of the weak instruments problem, although Hahn-Hausman (2002, 2003) propose a test that is useful in identifying when weak instruments is causing a problem. The source of the problem of small reported standard

---

[7] More generally, since the bias in the OLS estimate when EIV exists depends on the variance of the measurement error, or alternatively the R^2 of the regression, typically no bounds exist in the EIV problem for the estimated coefficient unless some judgment is made regarding the unknown variance. For further discussion, see Hausman (2001).

errors of the 2SLS estimator has not been discussed in the literature. Here we derive the source of the problem and offer a possible approach to fixing it.

The variance of 2SLS is derived in Theorem 1 and takes the usual form of $V_{2SLS} = \sigma_{\varepsilon\varepsilon}\Theta^{-1}$ where $\Theta = \pi'z'z\pi/n$ is assumed to be fixed. Now $\hat{\Theta}$ is not difficult to estimate since unbiased estimated of π follow from OLS on equation (1.2). Thus, the downward bias in the estimated 2SLS standard errors must arise from a downward biased estimate of $\sigma_{\varepsilon\varepsilon}$. We now derive the bias. The intuition follows from the fact that 2SLS is biased towards the OLS estimator, which minimizes $\hat{\sigma}_{\varepsilon\varepsilon}$. Thus, we find that the bias of the 2SLS estimator of β creates a bias in the 2SLS estimate of $\sigma_{\varepsilon\varepsilon}$. We find the bias to be:

Theorem 7: $E[\hat{\sigma}^2{}_{2SLS}] \approx \sigma_{\varepsilon\varepsilon} - \dfrac{2}{n}\dfrac{(K-2)\sigma^2{}_{\varepsilon v}}{\Theta} - \dfrac{1}{n}\sigma^2{}_{\varepsilon v} + \dfrac{1}{n}\dfrac{\sigma_{\varepsilon\varepsilon}\sigma_{vv}}{\Theta}$

Note that the leading term in the bias calculation of Theorem 5 is 2 times the bias of the 2SLS estimator from equation (2.1). As either the number of instruments grows or the covariance between the structural and reduced term stochastic disturbances becomes large, the bias in the estimation of $\sigma_{\varepsilon\varepsilon}$ will also become large. We now apply the normalization that we used above to find:

$$(5.1) \quad \begin{aligned} E[\hat{\sigma}^2{}_{2SLS}] &\approx 1 - \frac{2}{n}\frac{(K-2)\rho^2(1-R^2)}{R^2} - \frac{1}{n}\rho^2 + \frac{1}{n}\frac{(1-R^2)}{R^2} \\ &= 1 - \frac{1}{n}\frac{[(2K-4)\rho^2 - 1](1-R^2)}{R^2} - \frac{1}{n}\rho^2 \end{aligned}$$

The bias can be quite substantial as demonstrated by equation (5.1). The final term in equation (4.2) will typically be small so that it can be ignored. Equation (5.1) demonstrates that the downward bias can be substantial; in Monte-Carlo results we find that for $R^2 = .01$ and $\rho = 0.9$ that the mean bias of the 2SLS estimate of the variance varies from –70% to –80% as K, the number of instruments, increases from 5 to 30.[8] Thus, we note that the bias in the estimation even when K = 5 can be quite large. This finding explains the result that when weak instruments are present, the estimated standard

---

[8] The Monte-Carlo design is the same as in Hahn-Hausman (2002a).

errors of 2SLS can appear to be near those of OLS and small enough to allow the researcher to make conclusions about the likely true parameter value. However, with weak instruments these conclusions could be erroneous because of the substantial bias in the estimated standard error of the 2SLS estimator. Kleibergen (2002) also proposes an alternative approach to modify inferential procedures, but his approach is based on the LIML estimator rather than the 2SLS estimator. Hahn-Hausman (2003) and Hahn, Hausman, and Kuiersteiner (2003) discuss problems that may arise with this approach because of non-existence of moments of the LIML estimator.

We now consider the finding that the often used test of over identifying restrictions (OID test) rejects "too often" when weak instruments are present, i.e. the actual size of the test is considerably larger than the nominal size. See Hahn-Hausman (2002a), Table xx where the nominal size is 0.05 while the actual size is sometimes greater than 0.5. The OID test can be quite important since it tests the economic theory embodied in the model as discuss by e.g. Hausman (1983). In the weak instrument situation it may have increased importance given the substantial bias in the 2SLS estimator and the large MSE that we calculation in equations (3.1), (3.3) and (3.4). From Hausman (1983) we write the OID test as:

$$(5.2) \quad W = \frac{\hat{\varepsilon}' P_Z \hat{\varepsilon}}{\hat{\sigma}_{\varepsilon\varepsilon}}$$

W is distributed as chi-square with K-1 degrees of freedom. From equation (5.2), we see that a downward biased of $\sigma_{\varepsilon\varepsilon}$ can lead to substantial over-rejection and an upward biased size of the OID test. Thus, correcting for this problem can have an important effect on test results.

VI   Conclusions

We derive second order approximations for the bias and MSE of 2SLS (and the Nagar estimator) with both valid and invalid instruments. The derivation for invalid instruments is new, to the best of our knowledge. We find that substantial finite sample bias can occur when weak instruments exist which arises when the R^2 of the reduced

18

form regression is low, the number of instruments is high, or the correlation between the structural and reduced form stochastic terms ρ is high.

We then compare the bias and MSE of 2SLS with OLS. The OLS estimator is biased and inconsistent, but its smaller variance may make it preferable to 2SLS in a weak instruments situation. We determine straightforward and easily checked conditions under which 2SLS has smaller bias than OLS. These bias conditions carry over, in large part, to the MSE comparisons because changes in the bias term are quite important in changes in the MSE term given typical sample sizes of n=100 or larger. We find that for $R^2 \geq 0.1$, 2SLS is generally the preferred estimator. However, the econometrician can use our formulae to check the expected performance of 2SLS and OLS in a given situation given some a priori knowledge about likely parameter values.

We also demonstrate that a substantial bias exists in the 2SLS estimator for the variance of the stochastic disturbance, which lead to downward biased 2SLS standard errors and over-rejection of the test of over-identifying restrictions. We derive a formula for the bias that would allow for correction of the bias.

References

Angrist, J. and A. Krueger (1991): "Does Compulsory School Attendance Affect Schooling and Earnings," *Quarterly Journal of Economics,"* 106, 979-1014.

Card, D. (2001), "Estimating the Return to Schooling," *Econometrica*, 1127-1152

Griliches, Z. (1957), "Specification Bias in Estimates of Production Function," *Journal of Farm Economics*, 38, 8-20

Griliches, Z. (1997), "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica, 45, 1-22.*

Hahn, J., and J. Hausman (2002a): "A New Specification Test for the Validity of Instrumental Variables," *Econometrica*, 70, 163-189.

Hahn, J., and J. Hausman (2002b): "Notes on Bias in Estimators for Simultaneous Equation Models", *Economics Letters*, 75, 237-241.

Hahn, J., and J. Hausman (2003): "Weak Instruments: Diagnosis and Cures in Empirical Econometrics", *American Economic Review*

Hahn, J., J. Hausman, and G. Kuiersteiner (2003): "Estimation with Weak Instruments: Accuracy of Higher Order Bias and MSE Approximations," revised MIT mimeo

Hausman, J.A. (1978): "Specification Tests in Econometrics," *Econometrica*, 46, 1251 – 1271.

Hausman, J.A. (1983): "Specification and Estimation of Simultaneous Equation Models," in Z. Griliches and M. Intriligator, *eds., Handbook of Econometrics*, Vol. 1, Amsterdam: North Holland.

Hausman, J. (2001), "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left", *Journal of Economic Perspectives*, 2001

Imbens, G. (2003), "Sensivity to Exogeneity Assumptions in Program Evaluation," *American Economic Review*

Kleibergen, F. (2002): "Pivotal Statistics for Testing Structural Parameters in IV Regression," *Econometrica*, 70, 1781-1803

Manski, C. (1990): "Nonparametric Bounds on Treatment Effects," *American Economic Review Papers and Proceedings*, 80, 319-323.

Manski, C. (2003), *Partial Identification of Probability Distributions*, New York: Springer-Verlag.

Manski, C. and R. Pepper (2000), "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica*, 68, 997–1010.

Rosenbaum, P., and D. Rubin (1983), "Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome", *Journal of the Royal Statistical Society, Series B*, 45, 212-218.

Staiger, D., and J.H. Stock (1997): "IV Regression with Weak Instruments," *Econometrica*, 65, 557-586.

Stock. J.H., J. Wright and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in GMM," *Journal of Business and Economic Statistics*, 20, 518-529.

Figure 1
Critical Values for Rho
n=100 and K=5, 10, 30

rr
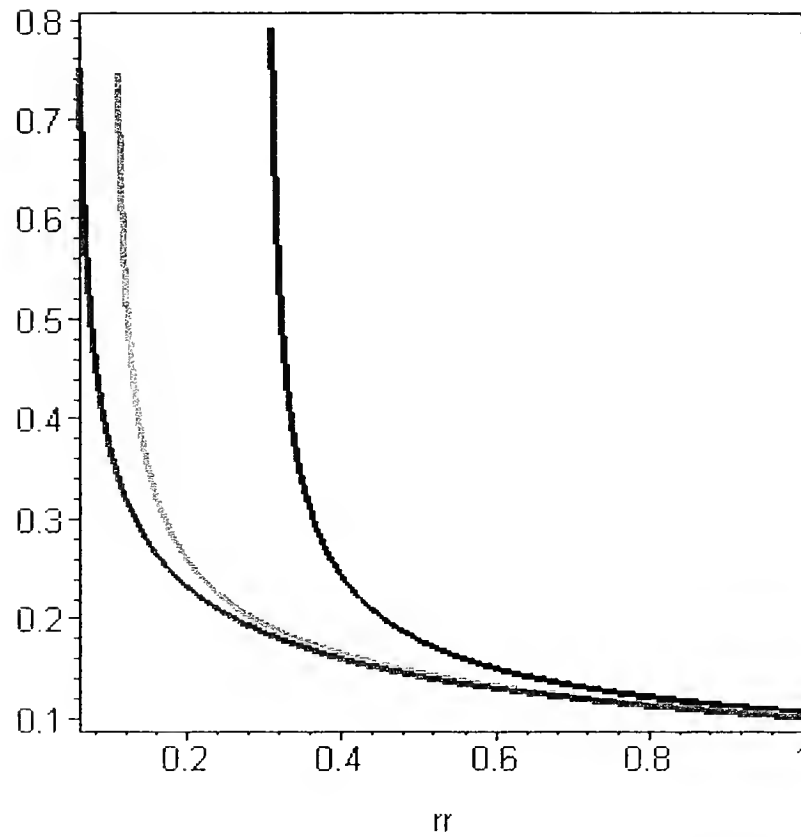
Table 1: Critical Values of ρ

R^2	0.01	0.1	0.2	0.3	0.5	0.7	0.9
K=5							
100	**	0.3677	0.2323	0.1863	0.1432	0.1210	0.1070
500	**	0.1423	0.1002	0.0818	0.0634	0.0536	0.0473
1000	0.3654	0.1002	0.0708	0.0578	0.0448	0.0378	0.0334
K=10							
100	**	**	0.2601	0.1949	0.1455	0.1220	0.1075
500	**	0.1445	0.1006	0.0819	0.0634	0.0536	0.0473
1000	**	0.1006	0.0708	0.0578	0.0448	0.0378	0.0334
K=10							
100	**	**	**	**	0.1789	0.1339	0.1135
500	**	0.1771	0.1050	0.0834	0.0638	0.0538	0.0474
1000	**	0.1049	0.0716	0.0581	0.0448	0.0379	0.0334

** denotes no critical value of ρ less than 1.0 exists

Figure 2: Ratio of 2SLS Bias to OLS Bias with Invalid Instruments

N=100, K=5, $\alpha = 0.1$

Figure 3

Critical Values for Alpha
n=100 and K=5, 10, 30

Figure 4

MSE 2SLS/MSE OLS

Alpha=0.1 and K=5

Figure 5

MSE 2SLS/MSE OLS
Alpha=0.3 and K=5

Figure 6

MSE 2SLS/MSE OLS

Alpha=0.1 and K=5

Table 2: Bias of $\hat{\sigma}^2$

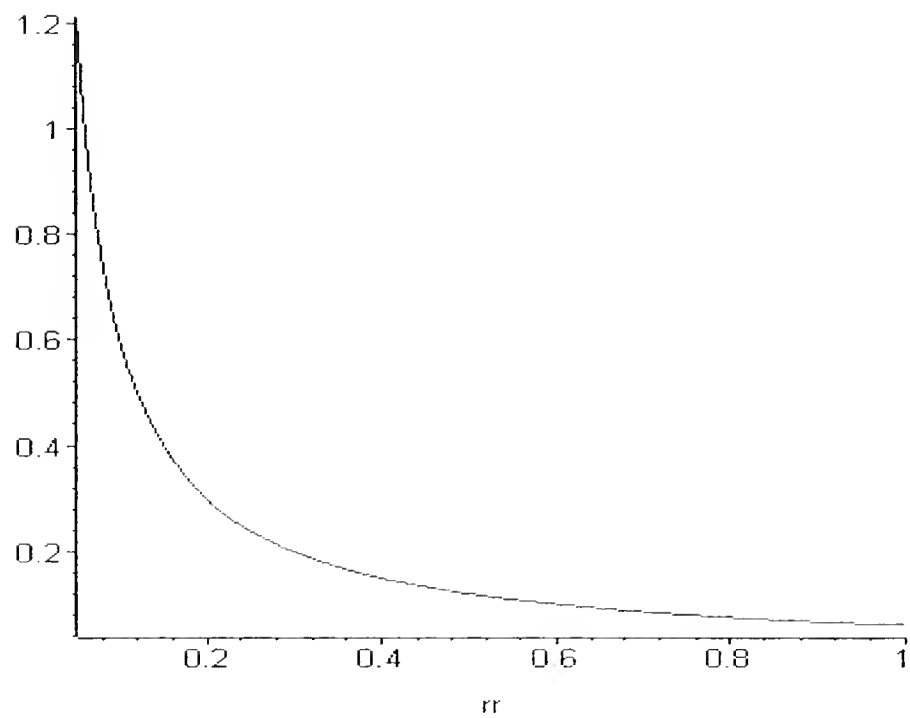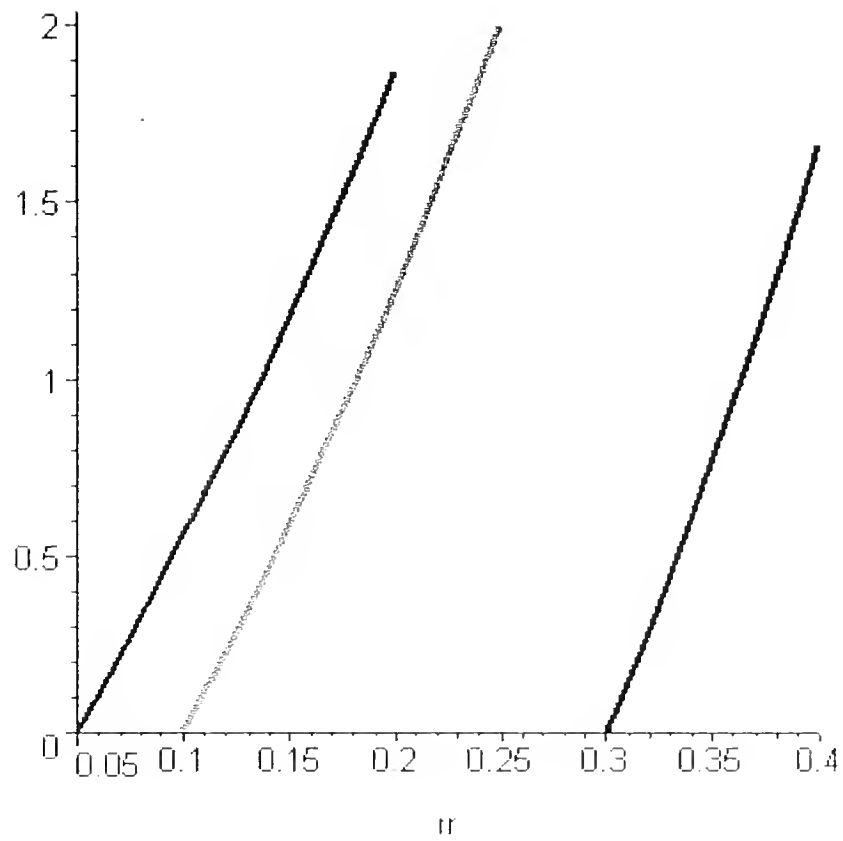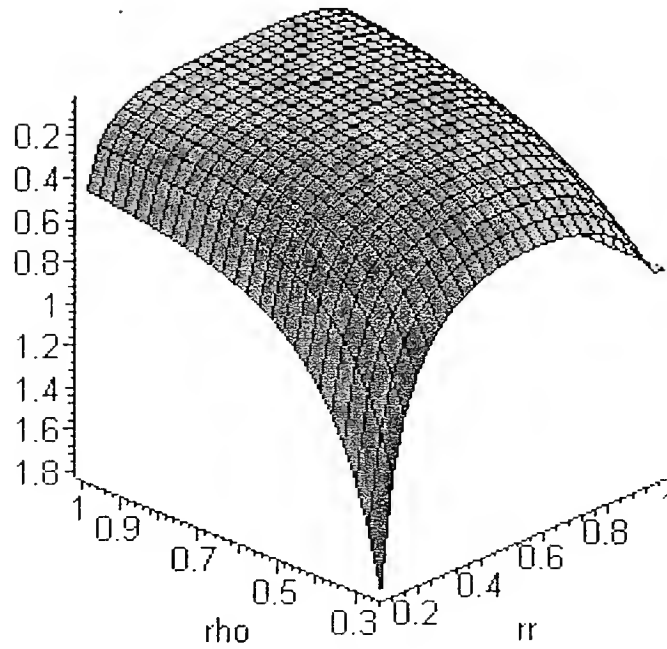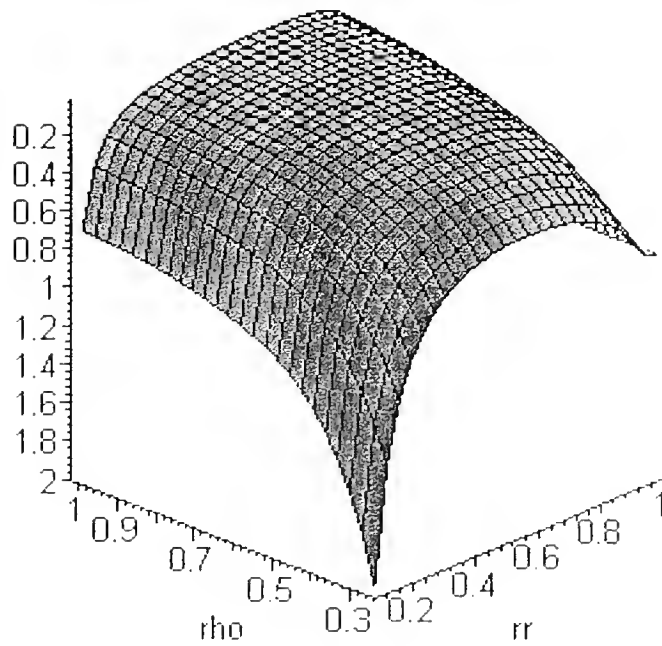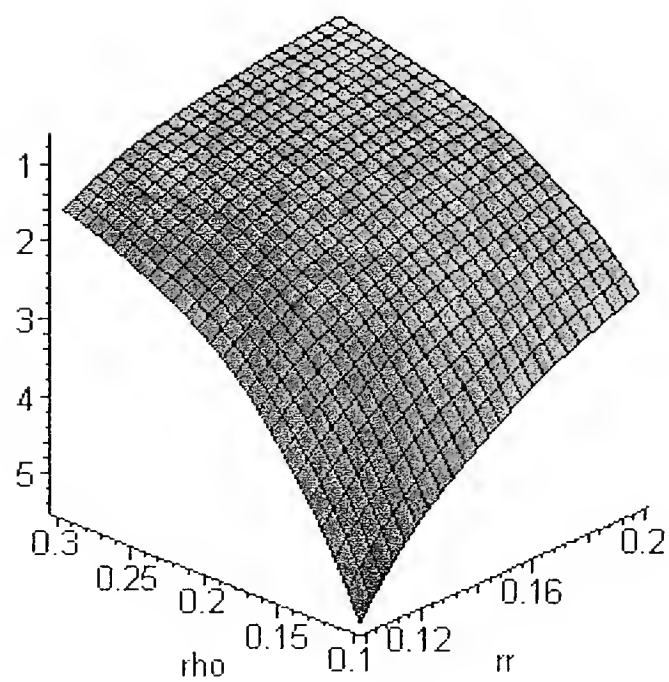| K	R^2	ρ	n=100 Mean Bias	Median Bias	MAE	RMSE	IQR	n=500 Mean Bias	Median Bias	MAE	RMSE	IQR	n=1000 Mean Bias	Median Bias	MAE	RMSE	IQR
5	0.01	0	0.25	0.10	0.31	0.81	0.33	0.14	0.07	0.16	0.34	0.16	0.09	0.04	0.10	0.21	0.11
5	0.01	0.5	-0.04	-0.16	0.28	0.56	0.26	-0.05	-0.15	0.22	0.35	0.22	-0.04	-0.12	0.20	0.33	0.23
5	0.01	0.9	-0.70	-0.76	0.72	0.73	0.12	-0.48	-0.58	0.55	0.61	0.28	-0.32	-0.41	0.43	0.51	0.34
10	0.01	0	0.09	0.06	0.17	0.25	0.25	0.08	0.05	0.10	0.17	0.13	0.06	0.04	0.07	0.12	0.09
10	0.01	0.5	-0.17	-0.20	0.21	0.25	0.19	-0.15	-0.19	0.19	0.22	0.13	-0.13	-0.18	0.18	0.21	0.14
10	0.01	0.9	-0.77	-0.78	0.77	0.77	0.06	-0.68	-0.71	0.68	0.69	0.13	-0.56	-0.59	0.56	0.58	0.18
30	0.01	0	0.01	0.00	0.12	0.15	0.20	0.03	0.02	0.06	0.08	0.10	0.03	0.02	0.05	0.06	0.07
30	0.01	0.5	-0.24	-0.24	0.24	0.26	0.15	-0.22	-0.23	0.22	0.23	0.08	-0.21	-0.22	0.21	0.22	0.07
30	0.01	0.9	-0.80	-0.80	0.80	0.80	0.04	-0.78	-0.78	0.78	0.78	0.03	-0.74	-0.75	0.74	0.75	0.06
5	0.1	0	0.06	0.03	0.15	0.23	0.22	0.02	0.01	0.06	0.07	0.09	0.01	0.01	0.04	0.05	0.06
5	0.1	0.5	-0.05	-0.11	0.21	0.30	0.27	-0.01	-0.03	0.11	0.14	0.18	-0.01	-0.02	0.08	0.10	0.13
5	0.1	0.9	-0.31	-0.39	0.41	0.47	0.34	-0.07	-0.11	0.19	0.24	0.28	-0.04	-0.06	0.14	0.17	0.21
10	0.1	0	0.04	0.03	0.13	0.18	0.21	0.01	0.01	0.05	0.07	0.09	0.01	0.01	0.04	0.05	0.06
10	0.1	0.5	-0.13	-0.16	0.19	0.23	0.21	-0.05	-0.06	0.11	0.14	0.16	-0.03	-0.04	0.08	0.10	0.12
10	0.1	0.9	-0.53	-0.56	0.53	0.55	0.19	-0.19	-0.21	0.23	0.26	0.23	-0.10	-0.12	0.15	0.18	0.19
30	0.1	0	0.00	0.00	0.12	0.15	0.20	0.01	0.01	0.05	0.07	0.09	0.01	0.01	0.04	0.05	0.06
30	0.1	0.5	-0.21	-0.22	0.22	0.24	0.16	-0.13	-0.14	0.14	0.16	0.11	-0.08	-0.09	0.10	0.11	0.10
30	0.1	0.9	-0.70	-0.70	0.70	0.70	0.07	-0.44	-0.45	0.44	0.45	0.12	-0.29	-0.30	0.29	0.31	0.13
5	0.3	0	0.01	0.00	0.12	0.15	0.20	0.00	0.00	0.05	0.06	0.09	0.00	0.00	0.04	0.04	0.06
5	0.3	0.5	-0.03	-0.06	0.16	0.20	0.25	0.00	-0.01	0.07	0.09	0.12	0.00	-0.01	0.05	0.06	0.09
5	0.3	0.9	-0.10	-0.15	0.24	0.29	0.32	-0.02	-0.03	0.11	0.13	0.17	-0.01	-0.02	0.08	0.10	0.12
10	0.3	0	0.01	0.00	0.12	0.15	0.20	0.00	0.00	0.05	0.06	0.09	0.00	0.00	0.04	0.04	0.06
10	0.3	0.5	-0.06	-0.08	0.15	0.19	0.22	-0.01	-0.02	0.07	0.09	0.12	-0.01	-0.01	0.05	0.07	0.09
10	0.3	0.9	-0.23	-0.26	0.27	0.31	0.25	-0.06	-0.06	0.11	0.14	0.17	-0.03	-0.04	0.08	0.10	0.12
30	0.3	0	0.00	-0.01	0.11	0.14	0.19	0.00	0.00	0.05	0.06	0.08	0.00	0.00	0.04	0.04	0.06
30	0.3	0.5	-0.15	-0.16	0.17	0.20	0.18	-0.05	-0.06	0.08	0.09	0.10	-0.03	-0.03	0.05	0.07	0.08
30	0.3	0.9	-0.47	-0.48	0.47	0.48	0.12	-0.17	-0.18	0.18	0.20	0.13	-0.09	-0.10	0.11	0.12	0.11

Note   Results are based on 5000 Monte Carlo runs True σ^2 was set equal to 1

A Bekker Asymptotic Distribution of 2SLS, OLS, and Nagar under Misspecification

Suppose that

$$
\begin{aligned}
y_{1i}^{\star} &= y_{2i}\beta + e_1 = \left(z_i'\pi_2\right)\beta + u_{1i} \\
y_{2i} &= z_i'\pi_2 + v_{2i}
\end{aligned}
$$

where

$$
\begin{pmatrix} u_{1i} \\ v_{2i} \end{pmatrix} \sim N\left(0, \begin{bmatrix} \omega_{1,1} & \omega_{1,2} \\ \omega_{1,2} & \omega_{2,2} \end{bmatrix}\right)
$$

Following is the Lemma reproduced from Hahn and Hausman (2001):

Lemma 1 *Let $U \equiv \begin{bmatrix} y_1^{\star} & y_2 \end{bmatrix}$. Assume that $\frac{K}{n} \to \alpha + o\left(n^{-1/2}\right)$, and that $\pi_2'z'z\pi_2/n$ is fixed at Θ. Let $\overline{S} \equiv U'P_zU$ and $S^{\perp} \equiv U'M_zU$. We then have*

$$
\sqrt{n}\left(\begin{pmatrix} n^{-1}\overline{S}_{11} \\ n^{-1}\overline{S}_{12} \\ n^{-1}\overline{S}_{22} \\ n^{-1}S_{11}^{\perp} \\ n^{-1}S_{12}^{\perp} \\ n^{-1}S_{22}^{\perp} \end{pmatrix} - \begin{pmatrix} \Theta \cdot \beta^2 + \frac{K}{n}\cdot\omega_{1,1} \\ \Theta \cdot \beta + \frac{K}{n}\cdot\omega_{1,2} \\ \Theta + \frac{K}{n}\cdot\omega_{2,2} \\ \left(1 - \frac{K}{n}\right)\cdot\omega_{1,1} \\ \left(1 - \frac{K}{n}\right)\cdot\omega_{1,2} \\ \left(1 - \frac{K}{n}\right)\cdot\omega_{2,2} \end{pmatrix}\right) \Rightarrow \mathcal{N}\left(0, \begin{bmatrix} \overline{\Lambda} & 0 \\ 0 & \Lambda^{\perp} \end{bmatrix}\right),
$$

where $\overline{\Lambda}$ and $\Lambda^{\perp}$ denote symmetric 3×3 matrices such that

$$
\begin{aligned}
\overline{\Lambda}_{1,1} &= 4\omega_{1,1}\Theta\beta^2 + 2\alpha\omega_{1,1}^2 \\
\overline{\Lambda}_{1,2} &= 2\omega_{1,1}\Theta\beta + 2\beta^2\Theta\omega_{1,2} + 2\alpha\omega_{1,1}\omega_{1,2} \\
\overline{\Lambda}_{1,3} &= 4\beta\Theta\omega_{1,2} + 2\alpha\omega_{1,2}^2 \\
\overline{\Lambda}_{2,2} &= \omega_{1,1}\Theta + \beta^2\Theta\omega_{2,2} + 2\Theta\omega_{1,2}\beta + \alpha\omega_{1,1}\omega_{2,2} + \alpha\omega_{1,2}^2 \\
\overline{\Lambda}_{2,3} &= 2\omega_{2,2}\Theta\beta + 2\Theta\omega_{1,2} + 2\alpha\omega_{2,2}\omega_{1,2} \\
\overline{\Lambda}_{3,3} &= 4\omega_{2,2}\Theta + 2\alpha\omega_{2,2}^2
\end{aligned}
$$

and

$$
\begin{aligned}
\Lambda_{1,1}^{\perp} &= 2\left(1-\alpha\right)\omega_{1,1}^2 \\
\Lambda_{1,2}^{\perp} &= 2\left(1-\alpha\right)\omega_{1,1}\omega_{1,2} \\
\Lambda_{1,3}^{\perp} &= 2\left(1-\alpha\right)\omega_{1,2}^2 \\
\Lambda_{2,2}^{\perp} &= \left(1-\alpha\right)\omega_{1,1}\omega_{2,2} + \left(1-\alpha\right)\omega_{1,2}^2 \\
\Lambda_{2,3}^{\perp} &= 2\left(1-\alpha\right)\omega_{2,2}\omega_{1,2} \\
\Lambda_{3,3}^{\perp} &= 2\left(1-\alpha\right)\omega_{2,2}^2
\end{aligned}
$$

1

Remark 1 *The ωs in the Lemma correspond to the "reduced form". It would be convenient to rewrite the above with structural form parameters. Because*

$$u = \varepsilon + \beta v$$

we can see that

$$
\begin{aligned}
\omega_{1,1} &= \sigma_{1,1} + 2\beta\sigma_{1,2} + \beta^2\sigma_{2,2} \\
\omega_{1,2} &= \sigma_{1,2} + \beta\sigma_{2,2} \\
\omega_{2,2} &= \sigma_{2,2}
\end{aligned}
$$

Lemma 2 *Suppose that $\frac{K}{\sqrt{n}} = \mu + o(1)$. Then we have*

$$
\sqrt{n}\left(\frac{y_2' P_z y_1^*}{y_2' P_z y_2} - \frac{\Theta \cdot \beta}{\Theta}\right) \Rightarrow N\left(\frac{\mu\sigma_{1,2}}{\Theta}, V_{2SLS}\right)
$$

$$
\sqrt{n}\left(\frac{y_2' P_z y_1^* - \frac{K}{n} y_2' M_z y_1^*}{y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2} - \beta\right) \Rightarrow N(0, V_{2SLS})
$$

$$
\sqrt{n}\left(\frac{y_2' y_1^*}{y_2' y_2} - \left(\beta + \frac{\sigma_{1,2}}{\Theta + \omega_{2,2}}\right)\right) \Rightarrow N(0, V_{OLS})
$$

Proof. Suppose that $\alpha = 0$. Using the previous Lemma, we obtain

$$
\sqrt{n}\left(\begin{pmatrix} n^{-1}\overline{S}_{12} \\ n^{-1}\overline{S}_{22} \end{pmatrix} - \begin{pmatrix} \Theta \cdot \beta + \frac{K}{n} \cdot \omega_{1,2} \\ \Theta + \frac{K}{n} \cdot \omega_{2,2} \end{pmatrix}\right) \Rightarrow N\left(0, \begin{bmatrix} (\beta^2\omega_{2,2} + 2\omega_{1,2}\beta + \omega_{1,1})\Theta & 2(\omega_{2,2}\beta + \omega_{1,2})\Theta \\ 2(\omega_{2,2}\beta + \omega_{1,2})\Theta & 4\omega_{2,2}\Theta \end{bmatrix}\right)
$$

$$
\sqrt{n}\left(\begin{pmatrix} n^{-1}(\overline{S}_{12} - \frac{K}{n}S_{12}^\perp) \\ n^{-1}(\overline{S}_{22} - \frac{K}{n}S_{22}^\perp) \end{pmatrix} - \begin{pmatrix} \Theta \cdot \beta \\ \Theta \end{pmatrix}\right) \Rightarrow N\left(0, \begin{bmatrix} (\beta^2\omega_{2,2} + 2\omega_{1,2}\beta + \omega_{1,1})\Theta & 2(\omega_{2,2}\beta + \omega_{1,2})\Theta \\ 2(\omega_{2,2}\beta + \omega_{1,2})\Theta & 4\omega_{2,2}\Theta \end{bmatrix}\right)
$$

and

$$
\sqrt{n}\left(\begin{pmatrix} n^{-1}(\overline{S}_{12} + S_{12}^\perp) \\ n^{-1}(\overline{S}_{22} + S_{22}^\perp) \end{pmatrix} - \begin{pmatrix} \Theta \cdot \beta + \omega_{1,2} \\ \Theta + \omega_{2,2} \end{pmatrix}\right)
$$

$$
\Rightarrow N\left(0, \begin{bmatrix} (\beta^2\omega_{2,2} + 2\omega_{1,2}\beta + \omega_{1,1})\Theta + \omega_{1,2}^2 + \omega_{1,1}\omega_{2,2} & 2(\omega_{2,2}\beta + \omega_{1,2})\Theta + 2\omega_{2,2}\omega_{1,2} \\ 2(\omega_{2,2}\beta + \omega_{1,2})\Theta + 2\omega_{2,2}\omega_{1,2} & 4\omega_{2,2}\Theta + 2\omega_{2,2}^2 \end{bmatrix}\right)
$$

Therefore, using Delta method, we obtain the following:

$$
\sqrt{n}\left(\frac{y_2' P_z y_1^*}{y_2' P_z y_2} - \frac{\Theta \cdot \beta + \frac{K}{n} \cdot \omega_{1,2}}{\Theta + \frac{K}{n} \cdot \omega_{2,2}}\right) \Rightarrow N\left(0, \frac{\sigma_{1,1}}{\Theta}\right)
$$

$$
\sqrt{n}\left(\frac{y_2' P_z y_1^* - \frac{K}{n} y_2' M_z y_1^*}{y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2} - \frac{\Theta \cdot \beta}{\Theta}\right) \Rightarrow N\left(0, \frac{\sigma_{1,1}}{\Theta}\right)
$$

$$
\sqrt{n}\left(\frac{y_2' y_1^*}{y_2' y_2} - \frac{\Theta \cdot \beta + \omega_{1,2}}{\Theta + \omega_{2,2}}\right) \Rightarrow N\left(0, \frac{\sigma_{1,1}}{(\Theta + \omega_{2,2})} - \frac{\sigma_{1,2}^2}{(\Theta + \omega_{2,2})^2} - 2\frac{\sigma_{1,2}^2\Theta^2}{(\Theta + \omega_{2,2})^4}\right)
$$

where we used the fact that

$$
\begin{aligned}
\omega_{1,1} &= \sigma_{1,1} + 2\beta\sigma_{1,2} + \beta^2\sigma_{2,2} \\
\omega_{1,2} &= \sigma_{1,2} + \beta\sigma_{2,2} \\
\omega_{2,2} &= \sigma_{2,2}
\end{aligned}
$$

Because $\frac{K}{\sqrt{n}} = \mu + o(1)$, we can see that

$$\sqrt{n}\left(\frac{\Theta \cdot \beta + \frac{K}{n} \cdot \omega_{1,2}}{\Theta + \frac{K}{n} \cdot \omega_{2,2}} - \frac{\Theta \cdot \beta}{\Theta}\right) = \frac{\frac{K}{\sqrt{n}}\sigma_{1,2}}{\Theta + \frac{1}{\sqrt{n}}\frac{K}{\sqrt{n}}\sigma_{2,2}} = \frac{\mu\sigma_{1,2}}{\Theta} + o(1)$$

and

$$\begin{aligned}
\sqrt{n}\left(\frac{y_2' P_z y_1^*}{y_2' P_z y_2} - \frac{\Theta \cdot \beta}{\Theta}\right) &=& \sqrt{n}\left(\frac{y_2' P_z y_1^*}{y_2' P_z y_2} - \frac{\Theta \cdot \beta + \frac{K}{n} \cdot \omega_{1,2}}{\Theta + \frac{K}{n} \cdot \omega_{2,2}}\right) + \sqrt{n}\left(\frac{\Theta \cdot \beta + \frac{K}{n} \cdot \omega_{1,2}}{\Theta + \frac{K}{n} \cdot \omega_{2,2}} - \frac{\Theta \cdot \beta}{\Theta}\right) \\
&=& \sqrt{n}\left(\frac{y_2' P_z y_1^*}{y_2' P_z y_2} - \frac{\Theta \cdot \beta + \frac{K}{n} \cdot \omega_{1,2}}{\Theta + \frac{K}{n} \cdot \omega_{2,2}}\right) + \frac{\mu\sigma_{1,2}}{\Theta} + o(1) \qquad \cdot \\
&\Rightarrow& N\left(\frac{\mu\sigma_{1,2}}{\Theta}, \frac{\sigma_{1,1}}{\Theta}\right)
\end{aligned}$$

∎

A.1 Asymptotic Distribution of 2SLS under Misspecification

Note that

$$\begin{aligned}
b_{2SLS} &=& \frac{y_2' P_z y_1}{y_2' P_z y_2} = \frac{y_2' P_z \left(y_1^* + \frac{1}{\sqrt{n}}z\gamma\right)}{y_2' P_z y_2} \\
&=& \frac{y_2' P_z y_1^*}{y_2' P_z y_2} + \frac{1}{\sqrt{n}}\frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1}y_2' P_z y_2}
\end{aligned}$$

But

$$\begin{aligned}
n^{-1}\left(z\pi_2 + v_2\right)' z\gamma &=& \Xi + n^{-1}v_2' z\gamma = \Xi + o_p(1) \\
n^{-1}y_2' P_z y_2 &=& n^{-1}\overline{S}_{22} = \Theta + o_p(1)
\end{aligned}$$

so that

$$\frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1}y_2' P_z y_2} = \frac{\Xi}{\Theta} + o_p(1)$$

It follows that

$$\begin{aligned}
\sqrt{n}\left(b_{2SLS} - \beta\right) &=& \sqrt{n}\left(\frac{y_2' P_z y_1^*}{y_2' P_z y_2} - \beta\right) + \frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1}y_2' P_z y_2} \\
&=& \sqrt{n}\left(\frac{y_2' P_z y_1^*}{y_2' P_z y_2} - \beta\right) + \frac{\Xi}{\Theta} + o_p(1) \\
&\Rightarrow& N\left(\frac{\Xi + \mu\sigma_{1,2}}{\Theta}, V_{2SLS}\right)
\end{aligned}$$

A.2 Asymptotic Distribution of OLS under Misspecification

Note that

$$\begin{aligned}
b_{OLS} &=& \frac{y_2' y_1}{y_2' y_2} = \frac{y_2'\left(y_1^* + \frac{1}{\sqrt{n}}z\gamma\right)}{y_2' y_2} \\
&=& \frac{y_2' y_1^*}{y_2' y_2} + \frac{1}{\sqrt{n}}\frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1}y_2' y_2}
\end{aligned}$$

3

But

$$n^{-1}\left(z\pi_2 + v_2\right)' z\gamma \;=\; \Xi + n^{-1} v_2' z\gamma = \Xi + o_p\left(1\right)$$
$$n^{-1} y_2' y_2 \;=\; n^{-1}\overline{S}_{22} + n^{-1} S_{22}^{\perp} = \Theta + \sigma_{2,2} + o_p\left(1\right)$$

so that

$$\frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1} y_2' P_z y_2} = \frac{\Xi}{\Theta + \sigma_{2,2}} + o_p\left(1\right)$$

It follows that

$$
\begin{aligned}
\sqrt{n}\left(b_{OLS} - \left(\beta + \frac{\sigma_{1,2}}{\Theta + \omega_{2,2}}\right)\right) \;&=\; \sqrt{n}\left(\frac{y_2' y_1^{\bullet}}{y_2' y_2} - \left(\beta + \frac{\sigma_{1,2}}{\Theta + \omega_{2,2}}\right)\right) + \frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1} y_2' y_2} \\
&=\; \sqrt{n}\left(\frac{y_2' y_1^{\bullet}}{y_2' y_2} - \left(\beta + \frac{\sigma_{1,2}}{\Theta + \omega_{2,2}}\right)\right) + \frac{\Xi}{\Theta + \sigma_{2,2}} + o_p\left(1\right) \\
\Rightarrow\; &N\left(\frac{\Xi}{\Theta + \sigma_{2,2}}, V_{OLS}\right)
\end{aligned}
$$

A.3 Asymptotic Distribution of Nagar under Misspecification

Note that

$$
\begin{aligned}
b_N \;&=\; \frac{y_2' P_z y_1 - \frac{K}{n} y_2' M_z y_1}{y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2} = \frac{y_2' P_z\left(y_1^{\bullet} + \frac{1}{\sqrt{n}} z\gamma\right) - \frac{K}{n} y_2' M_z\left(y_1^{\bullet} + \frac{1}{\sqrt{n}} z\gamma\right)}{y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2} \\
&=\; \frac{y_2' P_z y_1^{\bullet} - \frac{K}{n} y_2' M_z y_1^{\bullet}}{y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2} + \frac{1}{\sqrt{n}} \frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1}\left(y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2\right)}
\end{aligned}
$$

But

$$n^{-1}\left(z\pi_2 + v_2\right)' z\gamma \;=\; \Xi + o_p\left(1\right)$$
$$n^{-1}\left(y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2\right) \;=\; \Theta + o_p\left(1\right)$$

so that

$$\frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1}\left(y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2\right)} = \frac{\Xi}{\Theta} + o_p\left(1\right)$$

It follows that

$$
\begin{aligned}
\sqrt{n}\left(b_N - \beta\right) \;&=\; \sqrt{n}\left(\frac{y_2' P_z y_1^{\bullet} - \frac{K}{n} y_2' M_z y_1^{\bullet}}{y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2} - \beta\right) + \frac{n^{-1}\left(z\pi_2 + v_2\right)' z\gamma}{n^{-1}\left(y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2\right)} \\
&=\; \sqrt{n}\left(\frac{y_2' P_z y_1^{\bullet} - \frac{K}{n} y_2' M_z y_1^{\bullet}}{y_2' P_z y_2 - \frac{K}{n} y_2' M_z y_2} - \beta\right) + \frac{\Xi}{\Theta} + o_p\left(1\right) \\
\Rightarrow\; &N\left(\frac{\Xi}{\Theta}, V_{2SLS}\right)
\end{aligned}
$$

B Sensitivity Analysis

Consider a model with one endogenous regressor where other included exogenous variables are partialled out. The model takes the form where

$$y_i = x_i\beta + \varepsilon_i, \quad i = 1, \dots, n.$$

Denote the available instrument as z_i, and write the first stage regression as

$$x_i = z_i'\pi + v_i. \tag{1}$$

2SLS estimator is obviously given by

$$\widehat{\beta}_{2SLS} = \left[X'Z \left(Z'Z \right)^{-1} Z'X \right]^{-1} X'Z \left(Z'Z \right)^{-1} Z'Y,$$

where

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad Z = \begin{bmatrix} z_1' \\ \vdots \\ z_n' \end{bmatrix}.$$

What is the property of b if the exclusion restriction is in fact violated? In order to implement violation exclusion restriction, we <u>add</u> a little noise to ε_i, and consider a new model

$$y_i^* \left(\theta \right) = x_i\beta + z_i'\theta + \varepsilon_i^* \tag{2}$$

where

$$\varepsilon_i^* = z_i'\theta + \varepsilon_i$$

Let

$$\widehat{\beta}_{2SLS}^* \left(\theta \right) = \left[X'Z \left(Z'Z \right)^{-1} Z'X \right]^{-1} X'Z \left(Z'Z \right)^{-1} Z'Y^* \left(\theta \right)$$

and

$$b_{2SLS} \left(\theta \right) \equiv \text{plim}\, \widehat{\beta}_{2SLS}^* \left(\theta \right) - \beta$$

We would like to examine the maximal asymptotic bias $|b_{2SLS} \left(\theta \right)|$ for a small violation of exclusion restriction, i.e., the violation such that the correlation between $z_i'\theta$ and ε_i^* is some small number ψ. We argue that

$$\sqrt{\frac{n^{-1} \sum_{i=1}^n \widehat{\varepsilon}_i^2}{n^{-1} \sum_{i=1}^n x_i^2} \frac{1}{\widehat{\mathbb{R}}_f^2}} \sqrt{\frac{\psi^2}{1 - \psi^2}} \tag{3}$$

provides such measure of sensitivity. Here, $\widehat{\mathbb{R}}_f^2$ denotes the R^2 in the first stage.

5

B.1 Derivation of (3)

It can be shown that[1]

$$b_{2SLS}(\theta) = (\pi'\Phi\pi)^{-1}\pi'\Phi\theta$$

where

$$\Phi = \operatorname{plim} n^{-1}Z'Z$$

Note that

$$\frac{\partial b_{2SLS}(\theta)}{\partial\theta'} = (\pi'\Phi\pi)^{-1}\pi'\Phi$$

which is maximized when $\theta \propto \pi$. We therefore focus on the type of violation such that $\theta = \xi \cdot \pi$ for some scalar ξ.[2] Without loss of generality, we will write

$$b_{2SLS}(\theta) = b_{2SLS}(\xi \cdot \pi) = b_{2SLS}(\xi)$$

Note that the population R^2 in the regression of ε^* on z, which is equal to the square of the correlation ζ between ε_i^* and $z_i'\pi$, is equal to

$$\zeta^2 = \frac{\theta'\Phi\theta}{\theta'\Phi\theta + E\left[\varepsilon_i^2\right]} = \frac{\xi^2 \cdot \pi'\Phi\pi}{\xi^2 \cdot \pi'\Phi\pi + E\left[\varepsilon_i^2\right]} \tag{4}$$

and

$$b_{2SLS}(\xi) = (\pi'\Phi\pi)^{-1}\pi'\Phi(\xi \cdot \pi) = \xi \tag{5}$$

We can solve (4) for ξ^2, and obtain

$$\xi^2 = \frac{E\left[\varepsilon_i^2\right]}{\pi'\Phi\pi}\frac{\psi^2}{1 - \psi^2} \tag{6}$$

Now, note that the population R^2 in the first stage $\mathbb{R}_f^2$ is equal to

$$\mathbb{R}_f^2 = \frac{\pi'\Phi\pi}{E\left[x_i^2\right]}$$

which can be solved for $\pi'\Phi\pi$ as

$$\pi'\Phi\pi = \mathbb{R}_f^2 \cdot E\left[x_i^2\right] \tag{7}$$

---

[1] See next subsection for a slightly more general proof

[2] Maximization of $\|b_{2SLS}(\theta)\|^2$ with respect to θ fixing $\theta'\Phi\theta$ constant has the purpose of maximizing the asymptotic bias $b_{2SLS}(\theta)$ for a fixed population R^2 in the regression of ε^* on z. Because

$$|\pi'\Phi\theta|^2 \le (\pi'\Phi\pi) \cdot (\theta'\Phi\theta)$$

with equality when $\theta \propto \pi$, we can say that π is the direction that maximizes the sensitivity (inconsistency) of b_{2SLS} for a given amount of violation of exclusion restriction

Combining (6) and (7), we obtain

$$\xi^2 = \frac{E\left[\varepsilon_i^2\right]}{E\left[x_i^2\right]} \frac{1}{\mathbb{R}_f^2} \frac{\psi^2}{1 - \psi^2}$$

or

$$|\xi| = |b_{2SLS}(\xi)| = \sqrt{\frac{E\left[\varepsilon_i^2\right]}{E\left[x_i^2\right]} \frac{1}{\mathbb{R}_f^2}} \sqrt{\frac{\psi^2}{1 - \psi^2}} \tag{8}$$

We note that (8) can be approximated by the empirical counterpart

$$|\xi| = |b_{2SLS}(\xi)| \approx \sqrt{\frac{n^{-1}\sum_{i=1}^n \widehat{\varepsilon}_i^2}{n^{-1}\sum_{i=1}^n x_i^2} \frac{1}{\widehat{\mathbb{R}}_f^2}} \sqrt{\frac{\psi^2}{1 - \psi^2}} \tag{9}$$

B.2 Digression: Robustness of 2SLS

In general, we estimate β by

$$\widehat{\beta}_A = \left[(ZA)' X\right]^{-1} (ZA)' Y$$

and the counterpart under small misspecification is

$$\widehat{\beta}_A^*(\theta) = \left[(ZA)' X\right]^{-1} (ZA)' Y^*(\theta)$$

so that

$$
\begin{aligned}
b_A(\theta) &\equiv \operatorname{plim}\widehat{\beta}_A^*(\theta) - \beta \\
&= \operatorname{plim}\left[(ZA)' X\right]^{-1}(ZA)' Y^*(\theta) - \beta \\
&= \operatorname{plim}\left[(ZA)' X\right]^{-1}(ZA)'(X\beta + Z\theta + \varepsilon) - \beta \\
&= \beta + \operatorname{plim}\left[(ZA)' X\right]^{-1}(ZA)' Z\theta - \beta \\
&= \operatorname{plim}[A'Z'X]^{-1} A'Z'Z\theta \\
&= \operatorname{plim}\left[A'Z'Z(Z'Z)^{-1}Z'X\right]^{-1} A'Z'Z\theta \\
&= (A'\Phi\pi)^{-1} A'\Phi\theta
\end{aligned}
$$

Note that

$$\frac{\partial b_{2SLS}(\theta)}{\partial \theta'} = (\pi'\Phi\pi)^{-1}\pi'\Phi$$

and

$$\frac{\partial b_A(\theta)}{\partial \theta'} = (A'\Phi\pi)^{-1} A'\Phi$$

Instead of dealing with an awkward normalization involving the weight matrix Φ, it is convenient to use assume that $\Phi = I$. We then have

$$\frac{\partial b_{2SLS}(\theta)}{\partial \theta'} = (\pi'\pi)^{-1}\pi'$$

and

$$\frac{\partial b_A(\theta)}{\partial \theta'} = (A'\pi)^{-1} A'$$

Remark 2 *If there is only one instrument, then $\frac{\partial b_{2SLS}(\theta)}{\partial\theta'} = \pi^{-1}$. Therefore, small π indicates that 2SLS is sensitive to misspecification.*

Remark 3 *If there are multiple components in π, and if the first component of π is small relative to other components of π, then $\frac{\partial b_{2SLS}(\theta)}{\partial\theta_1}$ would be small, i.e., 2SLS is not very sensitive to the violation of the exclusion restriction in $z_{i,1}$.*

Remark 4 *Note that*

$$\left\|\frac{\partial b_{2SLS}(\theta)}{\partial\theta'}\right\|^2 = (\pi'\pi)^{-1} = \frac{1}{\|\pi\|^2}$$

and

$$\left\|\frac{\partial b_A(\theta)}{\partial\theta'}\right\|^2 = (A'\pi)^{-1}A'A(A'\pi)^{-1} = \frac{\|A\|^2}{(A'\pi)^2} \leq \frac{\|A\|^2}{\|A\|^2\|\pi\|^2} = \frac{1}{\|\pi\|^2} = \left\|\frac{\partial b_{2SLS}(\theta)}{\partial\theta'}\right\|^2$$

Therefore, 2SLS is the most robust estimator among the class of IV estimators b_A.

C   Higher Order Bias of $\widehat{\sigma}^2$

Our model is given by

$$y_i = x_i\beta + \varepsilon_i,$$

$$x_i = f_i + u_i = z_i'\pi + u_i, \qquad i = 1, \dots, n$$

where $(\varepsilon_i, u_i)'$ is homoscedastic and normal. We consider the 2SLS

$$\beta_{2SLS} \equiv \frac{x'Py}{x'Px}$$

and the related estimator for the variance of ε_i,

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\widehat{\beta}_{2SLS}\right)^2$$

We have the following characterization of $\widehat{\sigma}^2$

$$\begin{aligned}
\widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i - x_i\left(\widehat{\beta}_{2SLS} - \beta\right)\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 - 2\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i x_i\right)\left(\widehat{\beta}_{2SLS} - \beta\right) + \left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right)\left(\widehat{\beta}_{2SLS} - \beta\right)^2 \\
&= \frac{\varepsilon'\varepsilon}{n} \\
&\quad -2\left(\frac{\varepsilon'u}{n} + \frac{f'\varepsilon}{n}\right)\left(\widehat{\beta}_{2SLS} - \beta\right) \\
&\quad + \left(H + 2\frac{f'u}{n} + \frac{u'u}{n}\right)\left(\widehat{\beta}_{2SLS} - \beta\right)^2
\end{aligned}$$

where

$$H \equiv \frac{1}{n}f'f = \frac{1}{n}\pi'Z'Z\pi$$

8

Lemma 3

$$\sqrt{n}\left(\widehat{\beta}_{2SLS} - \beta\right) = \frac{1}{H}\sum_{j=1}^{7} T_j + o_p\left(\frac{1}{n}\right)$$

for

$$
\begin{aligned}
T_1 &= \frac{1}{\sqrt{n}}f'\varepsilon = O_p(1) \\[4pt]
T_2 &= \frac{u'P\varepsilon}{\sqrt{n}} = O_p\left(\frac{1}{\sqrt{n}}\right) \\[4pt]
T_3 &= -2\left(\frac{f'u}{n}\right)\frac{1}{H}\left(\frac{1}{\sqrt{n}}f'\varepsilon\right) = O_p\left(\frac{1}{\sqrt{n}}\right) \\[4pt]
T_4 &= 0 \\[4pt]
T_5 &= -\left(\frac{u'Pu}{n}\right)\frac{1}{H}\left(\frac{1}{\sqrt{n}}f'\varepsilon\right) = O_p\left(\frac{1}{n}\right) \\[4pt]
T_6 &= -2\left(\frac{f'u}{n}\right)\frac{1}{H}\left(\frac{u'P\varepsilon}{\sqrt{n}}\right) = O_p\left(\frac{1}{n}\right) \\[4pt]
T_7 &= 2^2\left(\frac{f'u}{n}\right)^2\frac{1}{H^2}\left(\frac{1}{\sqrt{n}}f'\varepsilon\right) = O_p\left(\frac{1}{n}\right)
\end{aligned}
$$

Proof. Note that 2SLS is a special case of the k-class estimator

$$\widehat{\beta}_k = \frac{x'Py - k \cdot x'My}{x'Px - k \cdot x'Mx}$$

for

$$k = \frac{a\theta + \frac{b}{n}}{1 - a\theta - \frac{b}{n}}$$

and θ is the "eigenvalue". Note that 2SLS corresponds to $a = 0$ and $b = 0$. The result follows from Donald and Newey (1998). ∎

We therefore obtain

Lemma 4

$$
\begin{aligned}
\widehat{\sigma}^2 &= \sigma_\varepsilon^2 \\
&+ \frac{1}{\sqrt{n}}\sqrt{n}\left(\frac{\varepsilon'\varepsilon}{n} - \sigma_\varepsilon^2\right) - \frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}T_1\right) \\
&- \frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}T_2 + \frac{1}{H}T_3\right) - \frac{2}{n}\sqrt{n}\left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right)\left(\frac{1}{H}T_1\right) - \frac{1}{n}\frac{T_1^2}{H} + \frac{1}{n}\frac{\sigma_u^2 T_1^2}{H^2} \\
&+ o_p\left(\frac{1}{n}\right)
\end{aligned}
$$

Proof. We have

$$
\begin{aligned}
\widehat{\sigma}^2 &= \frac{\varepsilon'\varepsilon}{n} \\
&\quad - \frac{2}{\sqrt{n}}\left(\frac{\varepsilon'u}{n} + \frac{f'\varepsilon}{n}\right)\left(\frac{1}{H}\sum_{j=1}^{7}T_j\right) \\
&\quad + \frac{1}{n}\left(H + 2\frac{f'u}{n} + \frac{u'u}{n}\right)\left(\frac{1}{H}\sum_{j=1}^{7}T_j\right)^2
\end{aligned}
$$

9

Because

$$T_1 = O_p(1)$$

$$T_2 = O_p\left(\frac{1}{\sqrt{n}}\right), \qquad T_3 = O_p\left(\frac{1}{\sqrt{n}}\right)$$

$$T_4 = 0, \qquad T_5 = O_p\left(\frac{1}{n}\right), \qquad T_6 = O_p\left(\frac{1}{n}\right), \qquad T_7 = O_p\left(\frac{1}{n}\right)$$

and

$$\frac{\varepsilon'u}{n} = O_p(1), \qquad \frac{f'\varepsilon}{n} - \frac{1}{\sqrt{n}}T_1 = O_p\left(\frac{1}{\sqrt{n}}\right)$$

$$\frac{f'u}{n} = O_p\left(\frac{1}{\sqrt{n}}\right), \qquad \frac{u'u}{n} = O_p(1)$$

we obtain

$$\widehat{\sigma}^2 = \frac{\varepsilon'\varepsilon}{n}$$

$$-\frac{2}{\sqrt{n}}\left(\frac{\varepsilon'u}{n}\right)\left(\frac{1}{H}\sum_{j=1}^{3}T_j\right) - \frac{2}{\sqrt{n}}\left(\frac{1}{\sqrt{n}}T_1\right)\left(\frac{1}{H}T_1\right)$$

$$+\frac{1}{n}\left(H + \frac{u'u}{n}\right)\left(\frac{1}{H}T_1\right)^2 + o_p\left(\frac{1}{n}\right)$$

Now, note that

$$\sqrt{n}\left(\frac{\varepsilon'\varepsilon}{n} - \sigma_\varepsilon^2\right) = O_p(1), \qquad \sqrt{n}\left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right) = O_p(1), \qquad \sqrt{n}\left(\frac{u'u}{n} - \sigma_u^2\right) = O_p(1)$$

We therefore obtain

$$\frac{\varepsilon'\varepsilon}{n} = \sigma_\varepsilon^2 + \frac{1}{\sqrt{n}}\sqrt{n}\left(\frac{\varepsilon'\varepsilon}{n} - \sigma_\varepsilon^2\right).$$

$$\frac{2}{\sqrt{n}}\left(\frac{\varepsilon'u}{n}\right)\left(\frac{1}{H}\sum_{j=1}^{3}T_j\right) = \frac{2}{\sqrt{n}}\left(\sigma_{\varepsilon u} + \left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right)\right)\left(\frac{1}{H}\sum_{j=1}^{3}T_j\right)$$

$$= \frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}\sum_{j=1}^{3}T_j\right) + \frac{2}{n}\sqrt{n}\left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right)\left(\frac{1}{H}T_1\right) + o_p\left(\frac{1}{n}\right),$$

and

$$\frac{1}{n}\left(H + \frac{u'u}{n}\right)\left(\frac{1}{H}T_1\right)^2 = \frac{1}{n}\frac{1}{H}T_1^2 + \frac{\sigma_u^2}{n}\frac{1}{H^2}T_1^2 + o_p\left(\frac{1}{n}\right)$$

It follows that

$$\widehat{\sigma}^2 = \sigma_\varepsilon^2 + \frac{1}{\sqrt{n}}\sqrt{n}\left(\frac{\varepsilon'\varepsilon}{n} - \sigma_\varepsilon^2\right)$$

$$-\frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}\sum_{j=1}^{3}T_j\right) - \frac{2}{n}\sqrt{n}\left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right)\left(\frac{1}{H}T_1\right) - \frac{2}{n}\frac{T_1^2}{H}$$

$$+\frac{1}{n}\frac{T_1^2}{H} + \frac{1}{n}\frac{\sigma_u^2 T_1^2}{H^2} + o_p\left(\frac{1}{n}\right)$$

or

$$
\begin{aligned}
\widehat{\sigma}^2 \; = \; & \sigma_\varepsilon^2 \\
& + \frac{1}{\sqrt{n}}\sqrt{n}\left(\frac{\varepsilon'\varepsilon}{n} - \sigma_\varepsilon^2\right) - \frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}T_1\right) \\
& - \frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}T_2 + \frac{1}{H}T_3\right) - \frac{2}{n}\sqrt{n}\left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right)\left(\frac{1}{H}T_1\right) - \frac{1}{n}\frac{T_1^2}{H} + \frac{1}{n}\frac{\sigma_u^2 T_1^2}{H^2} \\
& + o_p\left(\frac{1}{n}\right)
\end{aligned}
$$

∎

Condition 1 *Assume that we can ignore the $o_p\left(\frac{1}{n}\right)$ term in Lemma 4 in calculation of expectation.*

Theorem 1

$$
E\left[\widehat{\sigma}^2\right] \approx \sigma_\varepsilon^2 - \frac{2}{n}\frac{(K-2)\sigma_{\varepsilon u}^2}{H} - \frac{1}{n}\sigma_\varepsilon^2 + \frac{1}{n}\frac{\sigma_u^2 \sigma_\varepsilon^2}{H}
$$

where

$$
H \equiv \frac{1}{n}f'f = \frac{1}{n}\pi'Z'Z\pi
$$

Proof. From Lemma 4, we have

$$
\begin{aligned}
\widehat{\sigma}^2 \; = \; & \sigma_\varepsilon^2 \\
& + \frac{1}{\sqrt{n}}\sqrt{n}\left(\frac{\varepsilon'\varepsilon}{n} - \sigma_\varepsilon^2\right) - \frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}T_1\right) \\
& - \frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}T_2 + \frac{1}{H}T_3\right) - \frac{2}{n}\sqrt{n}\left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right)\left(\frac{1}{H}T_1\right) - \frac{1}{n}\frac{T_1^2}{H} + \frac{1}{n}\frac{\sigma_u^2 T_1^2}{H^2} \\
& + o_p\left(\frac{1}{n}\right)
\end{aligned}
$$

Because expected values of the $O_p\left(\frac{1}{\sqrt{n}}\right)$ terms in the second line are zero, it suffices to consider the $O_p\left(\frac{1}{n}\right)$ in the third line. First, we note that

$$
\begin{aligned}
E\left[T_2\right] &= E\left[\frac{u'P\varepsilon}{\sqrt{n}}\right] = \frac{1}{\sqrt{n}}K\sigma_{\varepsilon u} \\
E\left[T_3\right] &= E\left[-2\left(\frac{f'u}{n}\right)\frac{1}{H}\left(\frac{1}{\sqrt{n}}f'\varepsilon\right)\right] = -\frac{2}{\sqrt{n}}\sigma_{\varepsilon u}
\end{aligned}
$$

from which we obtain

$$
E\left[-\frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}T_2 + \frac{1}{H}T_3\right)\right] = -\frac{2}{n}\frac{(K-2)\sigma_{\varepsilon u}^2}{H}
$$

Second, we note that

$$
E\left[-\frac{2}{n}\sqrt{n}\left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right)\left(\frac{1}{H}T_1\right)\right] = 0
$$

due to symmetry. Third, we note that

$$
E\left[T_1^2\right] = H\sigma_\varepsilon^2
$$

11

from which we obtain

$$E\left[-\frac{1}{n}\frac{T_1^2}{H} + \frac{1}{n}\frac{\sigma_u^2 T_1^2}{H^2}\right] = -\frac{1}{n}\sigma_\varepsilon^2 + \frac{1}{n}\frac{\sigma_u^2\sigma_\varepsilon^2}{H}$$

We therefore obtain

$$\widetilde{E}\left[\widehat{\sigma}^2\right] = \sigma_\varepsilon^2 - \frac{2}{n}\frac{(K-2)\sigma_{\varepsilon u}^2}{H} - \frac{1}{n}\sigma_\varepsilon^2 + \frac{1}{n}\frac{\sigma_u^2\sigma_\varepsilon^2}{H}$$

■

.   .

Remark 5 *In order to understand Theorem 1, imagine a counter-factual situation where the first order asymptotic approximation for $\sqrt{n}\left(\widehat{\beta}_{2SLS} - \beta\right)$ is exact, i.e., write*

$$\sqrt{n}\left(\widehat{\beta}_{2SLS} - \beta\right) = \frac{1}{H}T_1$$

We would then have

$$\begin{aligned}\widehat{\sigma}^2 \quad &= \quad \sigma_\varepsilon^2 \\ &+ \frac{1}{\sqrt{n}}\sqrt{n}\left(\frac{\varepsilon'\varepsilon}{n} - \sigma_\varepsilon^2\right) - \frac{2}{\sqrt{n}}\sigma_{\varepsilon u}\left(\frac{1}{H}T_1\right) \\ &- \frac{2}{n}\sqrt{n}\left(\frac{\varepsilon'u}{n} - \sigma_{\varepsilon u}\right)\left(\frac{1}{H}T_1\right) - \frac{1}{n}\frac{T_1^2}{H} + \frac{1}{n}\frac{\sigma_u^2 T_1^2}{H^2} \\ &+ o_p\left(\frac{1}{n}\right)\end{aligned}$$

and

$$E\left[\widehat{\sigma}^2\right] \approx \sigma_\varepsilon^2 - \frac{1}{n}\sigma_\varepsilon^2 + \frac{1}{n}\frac{\sigma_u^2\sigma_\varepsilon^2}{H}$$

Therefore, Theorem 1 implies that the approximate mean of $\widehat{\sigma}^2$ is smaller by

$$\frac{2}{n}\frac{(K-2)\sigma_{\varepsilon u}^2}{H}$$

than would be expected out of first order asymptotic approximation.

Remark 6 *Theorem 1 can be understood from a different perspective. Note that the approximate bias of 2SLS is equal to*

$$\frac{1}{\sqrt{n}}E\left[T_2 + T_3\right] = \frac{(K-2)\sigma_{\varepsilon u}}{nH}$$

Roughly speaking, 2SLS is biased toward OLS, which minimizes $\frac{1}{n}\sum_{i=1}^n (y_i - x_i b)^2$ with respect to b. If the 2SLS $\widehat{\beta}_{2SLS}$ is close to the OLS $\widehat{\beta}_{OLS}$, then we should expect

$$\frac{1}{n}\sum_{i=1}^n\left(y_i - x_i\widehat{\beta}_{2SLS}\right)^2 \approx \frac{1}{n}\sum_{i=1}^n\left(y_i - x_i\widehat{\beta}_{OLS}\right)^2 \ll \frac{1}{n}\sum_{i=1}^n (y_i - x_i\beta)^2 = \frac{1}{n}\sum_i \varepsilon_i^2$$

Date Due